DARIAH-DE

Aufbau von Forschungsinfrastrukturen für die eHumanities

GEFÖRDERT VOM

Bundesministerium
für Bildung
und Forschung

Förderkennzeichen 01UG1110A bis M

Die in dieser Broschüre gewählte männliche Form bezieht immer gleichermaßen weibliche Personen ein. Auf eine Doppelbezeichnung wurde aufgrund besserer Lesbarkeit verzichtet.

Kontakt

Unter dariah-sub@sub.uni-goettingen.de werden Fragen zum Forschungsprojekt beantwortet. Weitere Informationen sind unter www.de.dariah.eu erhältlich.

Über das Projekt

Was ist DARIAH-DE?

DARIAH-DE unterstützt die mit digitalen Ressourcen und Methoden arbeitenden Geistes- und Kulturwissenschaftler in Forschung und Lehre. Das Projekt baut zu diesem Zweck eine Forschungsinfrastruktur für Werkzeuge und Forschungsdaten auf und entwickelt Materialien für Lehre und Weiterbildung im Bereich Digital Humanities (DH).

Steckbrief

DARIAH-DE ist der deutsche Beitrag der europaweit agierenden Forschungsinfrastruktur „DARIAH-EU – Digital Research Infrastructure for the Arts and Humanities".
www.dariah.eu

17 Kooperationspartner aus Geistes- und Kulturwissenschaften und Informationstechnologie, darunter Universitäten, Rechenzentren, fachspezifische Einrichtungen, eine Akademie der Wissenschaften, ein kommerzieller Partner sowie eine Bibliothek, bilden zurzeit das DARIAH-DE-Konsortium.

Projektkoordination: Niedersächsische Staats- und Universitätsbibliothek Göttingen.

Förderung: Bundesministerium für Bildung und Forschung (BMBF) Förderkennzeichen 01UG1110A bis M.

Laufzeit: Die Aufbauphase von DARIAH-EU ist zunächst bis 2016 geplant.

Über DARIAH-DE

Zentrale Aufgabe von DARIAH-DE ist es, die Interoperabilität von Werkzeugen und Forschungsdaten sowie deren Langzeitverfügbarkeit und Nachnutzung in den Digital Humanities zu ermöglichen. Hierbei orientiert sich DARIAH-DE an national sowie international gültigen und akzeptierten Standards und Vorgehensweisen und engagiert sich bei deren Weiterentwicklung. Darüber hinaus unterstützt und berät DARIAH-DE Fachwissenschaftler und Forschergruppen bei der Planung von geisteswissenschaftlichen Forschungsvorhaben in einer digitalen Umgebung.

Der kompetente Umgang mit digitalen Ressourcen, Konzepten und Methoden der Digital Humanities muss außerdem Eingang in die Lehre und Fortbildung von Geisteswissenschaftlern auf allen Stufen ihrer Ausbildung und beruflichen Praxis finden. Eng vernetzt mit den fachwissenschaftlichen Communities werden bestehende Studien- und Weiterbildungsangebote abgestimmt, besser sichtbar gemacht und weiterentwickelt. DARIAH-DE bietet darüber hinaus eigene Qualifizierungsmodule an, wie z. B. thematisch fokussierte, internationale Expertenworkshops.

Zur Etablierung einer digitalen Forschungskultur ist das Wissen um entsprechende Forschungsmethoden und Verfahren in den Geisteswissenschaften zu fördern. Der Einsatz dieser Methoden und Verfahren wird durch digitale Dienste und Werkzeuge unterstützt, die im Rahmen von DARIAH-DE konzipiert, entwickelt und als Basisinfrastruktur bereitgestellt werden. Um den Mehrwert von Methoden, Diensten und Werkzeugen der Digital Humanities zu veranschaulichen, werden die fachwissenschaftlichen Bedürfnisse anhand konkreter Forschungsfragen identifiziert. Basierend darauf werden einzelne Lösungen als sogenannte Demonstratoren entwickelt, um exemplarisch Potentiale und Methoden der Digital Humanities aufzuzeigen.

Forschungsdaten spielen eine zentrale Rolle im gesamten Forschungsprozess von der Recherche und Erfassung über die Analyse und Verarbeitung bis hin zur Publikation und anschließenden Nutzung auch durch Dritte. Eine wesentliche Grundlage dafür ist der vertrauensvolle und ungehinderte Zugriff auf diese Forschungsdaten und -ergebnisse. DARIAH-DE evaluiert, diskutiert und empfiehlt fachrelevante Standards für Daten, Metadaten, Lizenzen, Werkzeuge sowie Prozeduren und Organisationsstrukturen.

Die Forschungsinfrastruktur wird es Fachwissenschaftlern ermöglichen, in einer zunehmend digitalen Umgebung interdisziplinär, kollaborativ, nachhaltig und institutionenübergreifend zu arbeiten.

Zusammenspiel von DARIAH-EU und DARIAH-DE

DARIAH-EU ist eines von 48 Projekten auf der Roadmap des Europäischen Strategieforums für Forschungsinfrastrukturen (ESFRI). ESFRI wurde 2002 als strategisches Instrument von der Europäischen Kommission initiiert, um die Entwicklung der europäischen Forschungslandschaft voranzutreiben und diese im internationalen Wettbewerb zu stärken.

Aufgabe von ESFRI ist es, einen kohärenten und strategiegeleiteten Ansatz für Forschungsinfrastrukturen in Europa zu unterstützen sowie multilaterale Initiativen, die zur besseren Nutzung und Entwicklung von Forschungsinfrastrukturen führen, sowohl auf europäischer als auch auf globaler Ebene zu fördern.

Zu den Errungenschaften von ESFRI zählt die Entwicklung einer Rechtsform für Europäische Forschungsinfrastrukturkonsortien (ERIC), um den europäischen Forschungsverbünden eine stabile Organisations- und Rechtsform zu geben, die von allen EU-Mitgliedsstaaten anerkannt ist.

DARIAH-DE ist der deutsche Beitrag zu DARIAH-EU. Neben DARIAH-DE gibt es auch in weiteren europäischen Ländern nationale DARIAH-Aktivitäten, z. B. in Frankreich, den Niederlanden, Dänemark und Österreich – um nur einige zu nennen. Alle zusammengenommen fügen sie sich wie Puzzleteile zu DARIAH-EU, der europäischen vernetzten Forschungsinfrastruktur für die Geistes- und Kulturwissenschaften, zusammen.

Weitere Informationen zu ESFRI:
ec.europa.eu/research/esfri

Weitere Informationen zu DARIAH-EU:
www.dariah.eu

Was sind Digital Humanities?

Die Digital Humanities sind ein Forschungsgebiet am Schnittpunkt zwischen angewandter Informatik und den Geistes- und Kulturwissenschaften. In diesem interdisziplinären Kontext werden nicht-triviale computergestützte Methoden und Werkzeuge systematisch erforscht, entwickelt und angewandt, mit deren Hilfe neue geistes- und kulturwissenschaftliche Forschungsfragen gestellt und alte Forschungsfragen neu beantwortet werden können.

Die Forschungsmöglichkeiten werden durch Vernetzung von verteilten heterogenen digitalen Forschungsdaten sowie durch fachspezifische Methoden bereichert, wie z. B. digitales Edieren, quantitative Textanalyse, Visualisierung von komplexen Daten, interaktive historische Atlanten.

DHd-Blog – www.dhd-blog.org

Der Blog „Digital Humanities im deutschsprachigen Raum“ wurde federführend von DARIAH-DE und dem ebenfalls BMBF-geförderten TextGrid-Konsortium, das eine Virtuelle Forschungsumgebung für die textbasierten Digital Humanities errichtet, ins Leben gerufen. Er greift in Deutschland, Österreich und in der Schweiz die aktuellen Diskussionen in den Digital Humanities auf, informiert über neue Entwicklungen sowie Technologien und berichtet aus themenverwandten Forschungsvorhaben. Zudem fungiert er als Informationsbörse, indem er auf Förderprogramme, Veranstaltungen und Stellenausschreibungen im Feld der Digital Humanities aufmerksam macht.

Erklärtes Ziel ist es, möglichst alle in diesem Bereich aktiven Partner mit einzubeziehen, daher wird der DHd-Blog von DH-Experten aus Deutschland, Österreich und der Schweiz redaktionell betreut.

www.dhd-blog.org

Sie sind Digital-Humanities-Experte und möchten uns beim Bloggen unterstützen? Dann schreiben Sie eine E-Mail an:
dhdblog@mpiwg-berlin.mpg.de.

Affiliierte Projekte

Einige, z. T. auf europäischer Ebene agierende Projekte nutzen bereits die in DARIAH-DE eingerichtete Forschungsinfrastruktur. Eine Auflistung der Projekte finden Sie auf der DARIAH-DE Website:
www.de.dariah.eu

Sind Sie mit Ihrem Forschungsvorhaben, Ihrem Projekt oder Ihrer Forschergruppe an einer Zusammenarbeit mit DARIAH-DE interessiert? Dann nehmen Sie Kontakt mit uns auf:
dariah-sub@sub.uni-goettingen.de

DARIAH-DE: Forschung in den Digital Humanities

Virtuelle Forschungsumgebungen

Virtuelle Forschungsumgebungen (Virtual Research Environment, VRE) sind wesentliche Komponenten moderner Forschungsinfrastrukturen und dabei ein komplexes und sich ständig weiterentwickelndes Konzept. DARIAH-DE folgt der nachstehenden Definition von VREs der Allianz-Initiative:

„Unter einer virtuellen Forschungsumgebung wird hier eine Arbeitsplattform verstanden, die kooperative Forschungstätigkeiten durch mehrere Wissenschaftler an unterschiedlichen Orten zu gleicher Zeit ermöglicht. Inhaltlich unterstützt sie potentiell den gesamten Forschungsprozess – von der Erhebung, der Diskussion und weiteren Bearbeitung der Daten bis zur Publikation der Ergebnisse – während sie technologisch vor allem auf Softwarediensten und Kommunikationsnetzwerken basiert und Zugang zu realen Forschungsressourcen, wie z. B. Daten, Analyse- und Bearbeitungswerkzeuge, wissenschaftliche Geräte und Diensten gewährt.“

Virtuelle Forschungsumgebungen – Ein Leitfaden (Stand 30.09.2011), Arbeitsgruppe „Virtuelle Forschungsumgebungen“ in der Schwerpunktinitiative „Digitale Information“ der Allianz der deutschen Wissenschaftsorganisationen, S. 1, www.allianzinitiative.de/fileadmin/user_upload/Leitfaden_VRE_de.pdf

Modell-VRE

Die Geistes- und Kulturwissenschaften betrachten VREs als zunehmend wichtige Instrumente für ihre Forschung und folgen damit dem Beispiel der Naturwissenschaften, wo VREs längst Teil des Forschungsalltags sind. In DARIAH-DE wird auf der Basis des VRE-Leitfadens der Allianz-Initiative eine Modell-VRE mit Leitfaden für die Planung und Entwicklung eigener VREs erstellt, die speziell auf die Anforderungen der Geistes- und Kulturwissenschaften zugeschnitten ist.

Der Leitfaden beinhaltet u. a.:

1. eine Liste von Fragen zur Klärung bei ersten VRE-Planungsgesprächen zwischen Geistes- und Kulturwissenschaftlern, Software-Entwicklern und anderen Beteiligten,

2. die für fast alle VREs benötigten technischen Voraussetzungen, z. B. Rechte-Management,

3. einen Überblick über aktuelle Möglichkeiten der Drittmittelförderung,

4. Hinweise auf häufig auftretende Schwierigkeiten bei der Entwicklung von VREs sowie Lösungsvorschläge.

Nutzungsszenarien und -beispiele (Demonstratoren)

DARIAH-DE konzipiert und entwickelt in Kooperation mit Partnerinstitutionen fachspezifisch relevante Anwendungen für die Geistes- und Kulturwissenschaften, die als Demonstratoren bezeichnet werden.

Diese DARIAH-DE-Demonstratoren zeigen beispielhaft, welche Möglichkeiten die Digital Humanities eröffnen und wie die von DARIAH-DE bereitgestellte Forschungsinfrastruktur Fachwissenschaftler bei ihrer wissenschaftlichen Arbeit unterstützen kann.

Jeder DARIAH-DE-Demonstrator hat exemplarischen Charakter und bildet im prototypischen Betrieb nur einen Ausschnitt dessen ab, was mit digitalen Methoden und Verfahren bereits heute möglich ist. Um Forschungsprozesse nachbilden zu können, orientiert sich jeder Demonstrator an den folgenden vier Komponenten:

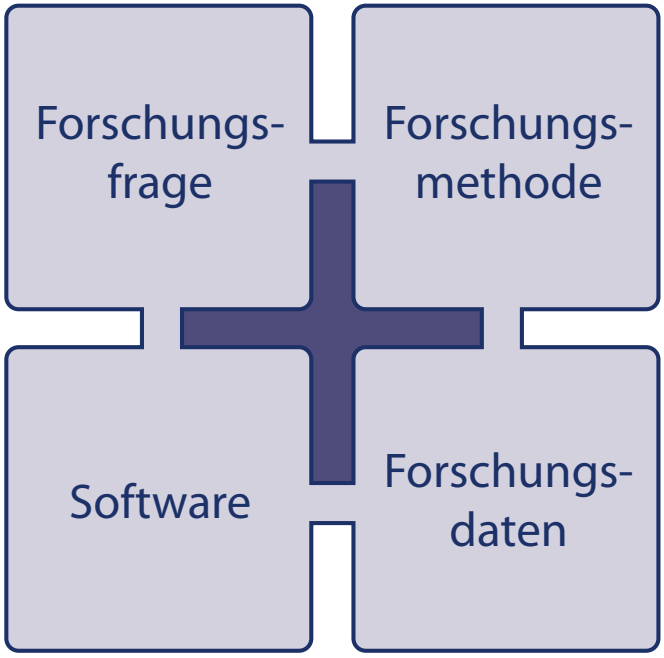


Abb. 1: Komponenten der DARIAH-DE-Demonstratoren.

Damit DARIAH-DE-Demonstratoren ihren Zweck effizient erfüllen können, wurde bei ihrer Konzeption besonderer Wert auf diese fünf essentiellen Kriterien gelegt:

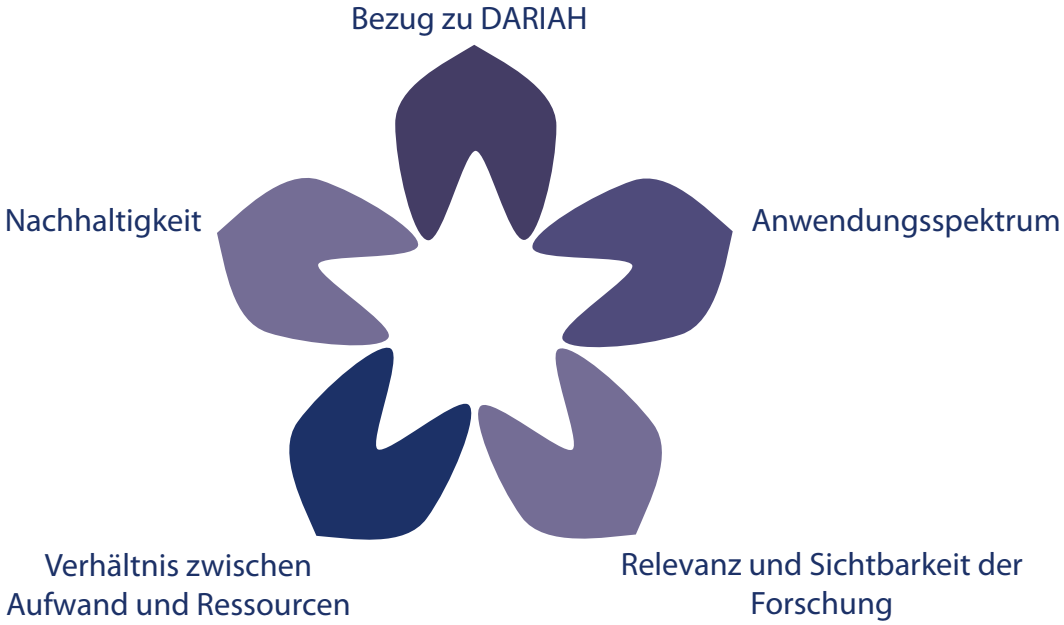


Abb. 2: Anforderungen an DARIAH-DE-Demonstratoren.

Folgende Demonstratoren werden zurzeit in DARIAH-DE entwickelt:

- **Archive-in-a-Box:**
  Empfehlungen und Konzepte für ein Archivsystem für geistes- und kulturwissenschaftliche Datenzentren und Forschungsverbünde.

- **Exploring TextGridRep with Voyant:**
  Verknüpfung der Voyant-Werkzeuge (Textanalyse) mit der digitalen Bibliothek des TextGrid-Repositories.

- **Generische Suche (vgl. Abb. 4):**
  Verteilte Metasuche über die in den DARIAH-DE-Registries hinterlegten digitalen Sammlungen.

- **Geo-Browser:**
  Nachnutzung der Europeana 4D-Schnittstelle zur Darstellung raumbezogener Daten in einem zeitlichen Kontext.

- **Interoperabilität durch Normdaten:**
  Semantische Verknüpfung von Forschungsdaten unterschiedlicher Provenienz durch die darin verwendeten Normdaten.

- **MEISE – Music Encoding Initiative Score Editor:**
  Spezieller Noteneditor zur Arbeit mit dem Standard MEI (Music Encoding Initiative) für die Erfassung schriftlich notierter Formen von Musik.

- **Monasterium Virtuelles Deutsches Urkundennetzwerk:**
  Virtuelle Forschungsumgebung für Urkunden des Mittelalters und der Frühen Neuzeit aus verschiedenen Beständen mehrerer deutscher Archive.

- **Personendaten-Repositorium:**
  Aufbau und freie Verfügbarkeit einer Architektur zur Vereinigung heterogener prosopographischer Datenbestände.

- **Virtuelles Skriptorium St. Matthias:**
  Virtuelle Rekonstruktion des noch erhaltenen Bestands der mittelalterlichen Bibliothek der Benediktinerabtei St. Matthias in Trier.

Recherche (DARIAH-Bibliographie)

Bei der DARIAH-Bibliographie, erreichbar unter www.zotero.org/groups/dariah_bibliography, handelt es sich um eine online zugängliche Auswahlbibliographie. Diese wird kollaborativ erstellt, regelmäßig aktualisiert sowie um neue Themen erweitert.

Die Bibliographie ist derzeit nach den folgenden Themen der digitalen Geistes- und Kulturwissenschaften geordnet: „Perspectives on Digital Humanities", „Text Analysis", „Virtual Research Environments", „Software and Services", „Digital Humanities Theses" und „Digital Scholarly Publishing". Da die Bibliographie auf Zotero basiert, können nützliche Funktionen mit dem kostenlosen Plug-in für Firefox oder in der eigenständigen zotero-Software genutzt werden. Für Hinweise zur Installation, zur Anmeldung und zu den Funktionen von Zotero siehe www.zotero.org/support.

www.zotero.org/groups/dariah_bibliography

Bei Fragen zur DH-Forschung in DARIAH-DE wenden Sie sich bitte an
forschung@dariah.eu

Zugang zu Forschungsdaten (Schema Registry, Collection Registry, Generische Suche)

Geistes- und Kulturwissenschaftler nutzen für ihre wissenschaftliche Tätigkeit unterschiedliche Medientypen: Bilder, handschriftliche Notizen, gedruckte Texte, Geo-Daten, archäologische Funde, Noten, Videos etc., die heute zunehmend in digitaler Form verfügbar sind. Dem geistes- und kulturwissenschaftlichen Erkenntnisinteresse, der unterschiedlichen Rezeption und Erforschung sind in methodischer und inhaltlicher Hinsicht nur noch wenige Grenzen gesetzt. Deshalb zählt der Zugang zu Forschungsdaten zu den wichtigsten Aspekten in der digitalen geistes- und kulturwissenschaftlichen Forschung. Dieser erweist sich für Forscher besonders dann als nur schwer überwindbare Hürde, wenn heterogene Daten nicht unter einer einheitlichen Oberfläche auffind- und durchsuchbar sind.

Schema Registry und Collection Registry

Als einer der zentralen Bestandteile der DARIAH-DE-Forschungsinfrastruktur wird eine Schema Registry entwickelt. Die Interoperabilität zwischen Forschungsdaten unterschiedlicher institutioneller und disziplinärer Herkunft wird dabei durch Mappings und Metadaten-Crosswalks ermöglicht.

Bei einem Mapping werden Relationen zwischen zwei verschiedenen Metadatenschemata gebildet. Ein Crosswalk ist die Überführung der Daten, die nach einem Schema modelliert sind, in eine andere Datenstruktur, basierend auf der Grundlage des Mappings. Die Quelldaten verbleiben selbstverständlich unverändert dezentral in den einzelnen Sammlungen. Diese sind in der Collection Registry verzeichnet, die ein weiterer zentraler Bestandteil der DARIAH-DE-Forschungsinfrastruktur ist.

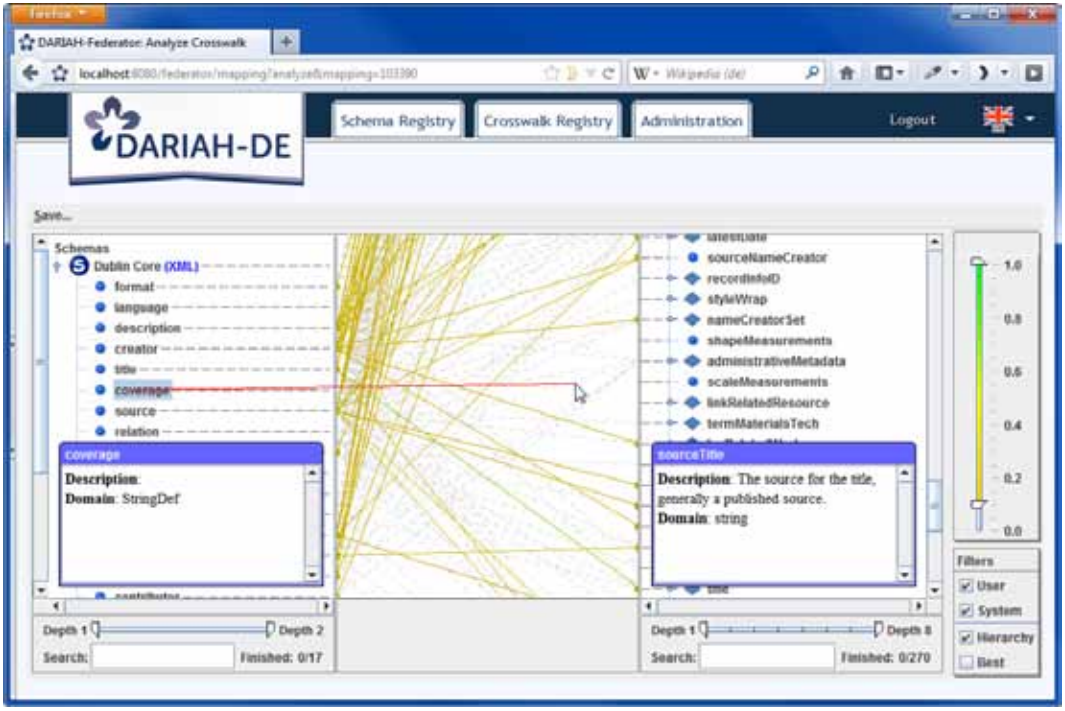


Abb. 3: Crosswalk innerhalb der Schema Registry.

Generische Suche

Für das Auffinden von heterogenen verteilten digitalen Forschungsdaten entwickelt DARIAH-DE eine generische Suchfunktionalität. Sie erlaubt die Abstraktion von der Heterogenität auf Sammlungs- und Datenmodellebene und nutzt semantische Zusammenhänge zwischen verschiedenen Sammlungen durch die Anwendung der Crosswalks.

Das generische Konzept der Suche besteht darin, dass Anfragen sowohl über eine große Zahl von Sammlungen (Breitensuche) als auch mit Fokus auf bestehende Zusammenhänge in strukturierten und semi-strukturierten Daten (Tiefensuche) ermöglicht werden. Die Anzahl möglicher Suchfacetten sinkt, wenn die Crosswalks nur wenige Übereinstimmungen aufweisen. Sind zwischen Sammlungen gar keine Crosswalks möglich, wird die Generische Suche als einfache feldübergreifende Volltextsuche realisiert.

Die Suchoberfläche (siehe Abb. 4) interagiert für die Vorbereitung der Suche mit der Collection Registry sowie mit der Schema Registry. Die Suchparameter werden durch den Benutzer festgelegt. Für die Interpretation von Schemata sowie die Auflösung von Referenzen in den Forschungs- und Metadaten greift die Generische Suche auch auf externe Quellen zu, wie z. B. den Getty Thesaurus für geographische Normdaten oder die Gemeinsame Normdatei (GND) für persistente Identifikatoren von Personennamen.

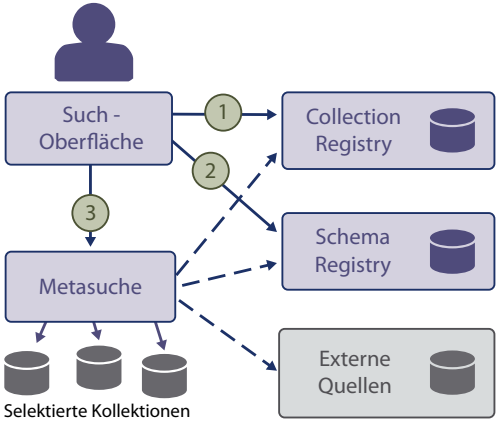


Abb. 4: Generische Suche.

Bei Fragen rund um Forschungsdaten in DARIAH-DE wenden Sie sich bitte an
forschungsdaten@dariah.eu

DARIAH-DE-Empfehlungen für Forscher

Umgang mit Forschungsdaten

Interoperabilität, Nachnutzbarkeit, Langlebigkeit der Daten sowie Vertrauen in ihre Qualität sind die zentralen Themen zwischen fächerübergreifender Perspektive und fachlicher Spezifik. Der nachhaltige Umgang mit Forschungsdaten ermöglicht hierbei die Nachnutzung der Forschungsdaten über fachliche, institutionelle und nationale Grenzen hinweg. Die Langzeitarchivierung und -verfügbarkeit sowie Interpretierbarkeit von Forschungsdaten und Metadaten (u.a. zur möglichst exakten Beschreibung der Forschungsdaten) ist dabei Grundvoraussetzung für eine funktionierende Forschungsinfrastruktur.

Anwendung von Standards

DARIAH-DE bietet weitgefächerte Angebote, um Wissenschaftlern den Einstieg in ein Forschungsprojekt und insbesondere dessen digitales Datenmanagement zu erleichtern.

Die Einhaltung und Weiterentwicklung von Standards und die Einrichtung einer (automatisierten) Qualitätskontrolle – etwa durch Metadatenvalidierung und eine transparente Versionsverwaltung – garantieren die Verlässlichkeit und Vertrauenswürdigkeit der registrierten Forschungsdaten.

In DARIAH-DE stehen u.a. die folgenden disziplinären und interdisziplinären Forschungsdaten- und Metadatenstandards im Fokus:

- **ADeX – Archaeological Data eXchange:**
  Standard für den Austausch von archäologischen Fachdaten

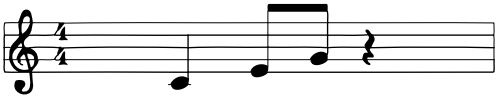
- **CEI – Charters Encoding Initiative:**
  Standard zur Kodierung von historischen Urkunden

- **TEI – Text Encoding Initiative:**
  Standard zur strukturierten Erfassung textbezogener Daten

- **EpiDoc – Epigraphic Documents:**
  Standard für die Wiedergabe von Inschriften in TEI XML

- **CIDOC-CRM – CIDOC Conceptual Reference Model:**
  Ontologie für Informationen zum Kulturerbe

- **MEI – Music Encoding Initiative:**
  Standard zur Erfassung notierter Formen von Musik



```
<scoreDef>
    <staffGrp>
        <staffDef label="Violin" n="0" key.sig="0" meter.count="4"
            meter.unit="4" clef.shape="G"/>
    </staffGrp>
</scoreDef>
<section>
    <measure n="1">
        <staff>
            <layer>
                <note pname="c" oct="4" dur="4" stem.dir="up"/>
                <beam>
                    <note pname="e" oct="4" dur="8" stem.dir="up"/>
                    <note pname="g" oct="4" dur="8" stem.dir="up"/>
                </beam>
                <rest dur="8"/>
            </layer>
        </staff>
    </measure>
</section>
```

Abb. 5: Ein in MEI kodiertes Notenbeispiel mit korrespondierendem Notenbild.

Die Anwendung solcher international etablierter Standards ist ein effektiver und dringlich zu empfehlender Weg, um Interoperabilität, Konsistenz und Nachhaltigkeit der erstellten Daten zu gewährleisten.

Lizenzierung und rechtliche Fragen

Im Umgang mit Forschungsdaten sind neben den technischen Herausforderungen auch die rechtlichen Rahmenbedingungen zu betrachten, die sich aufgrund von disziplinspezifischen Anforderungen und dem jeweiligen Anwendungsfall ergeben. In DARIAH-DE werden daher Datenlizenzen geprüft und lizenzrechtliche Fragestellungen im Hinblick auf Nachhaltigkeit und Interoperabilität untersucht, um Empfehlungen für den Umgang mit ihnen auszusprechen.

Das eigene Repository (Archive-in-a-Box)

Forscher verfügen häufig über Forschungsdaten, die sie kontrolliert und nachhaltig archivieren möchten. Archive-in-a-Box bietet ein Konzept, um Forschungsdokumente in einer digitalen Datensammlung – einem sogenannten Datenrepositorium – durch den Aufbau lokaler technischer Archivierungssysteme für Geistes- und Kulturwissenschaftler, Forschungszentren und Forschungsverbünde zu sichern und zu verwalten.

Ein Leitfaden informiert über lokal installierbare und gebührenfreie Software-Pakete, die die Anschlussfähigkeit an aktuelle und mögliche künftige technische Entwicklungen sicherstellen.

Die Software-Pakete enthalten eine umfangreiche Dokumentation, sind ressourcenschonend und berücksichtigen die gängigen Standards, Daten- und Metadatenformate (TEI, Dublin Core, OAI-PMH u. a.).

Folgt man diesen DARIAH-DE-Empfehlungen, sind die archivierten Dokumente und Daten kollaborativ bearbeitbar und intuitiv zu verwalten. Ein weiterer Vorteil ist das schnelle und unkomplizierte Einbringen von Daten im optimalen Zusammenspiel mit der DARIAH-DE-Forschungsinfrastruktur, die weitere Optionen einschließt:

- Nutzerverwaltung und Anmeldung mittels zuverlässiger Authentifizierung,
- eindeutige elektronische Kennzeichnung einzelner Dokumente zur dauerhaften Auffindbarkeit,
- Rückgriff auf DARIAH-DE-Suchfunktionen sowie auf Forschungsdatenschemata u. v. m.

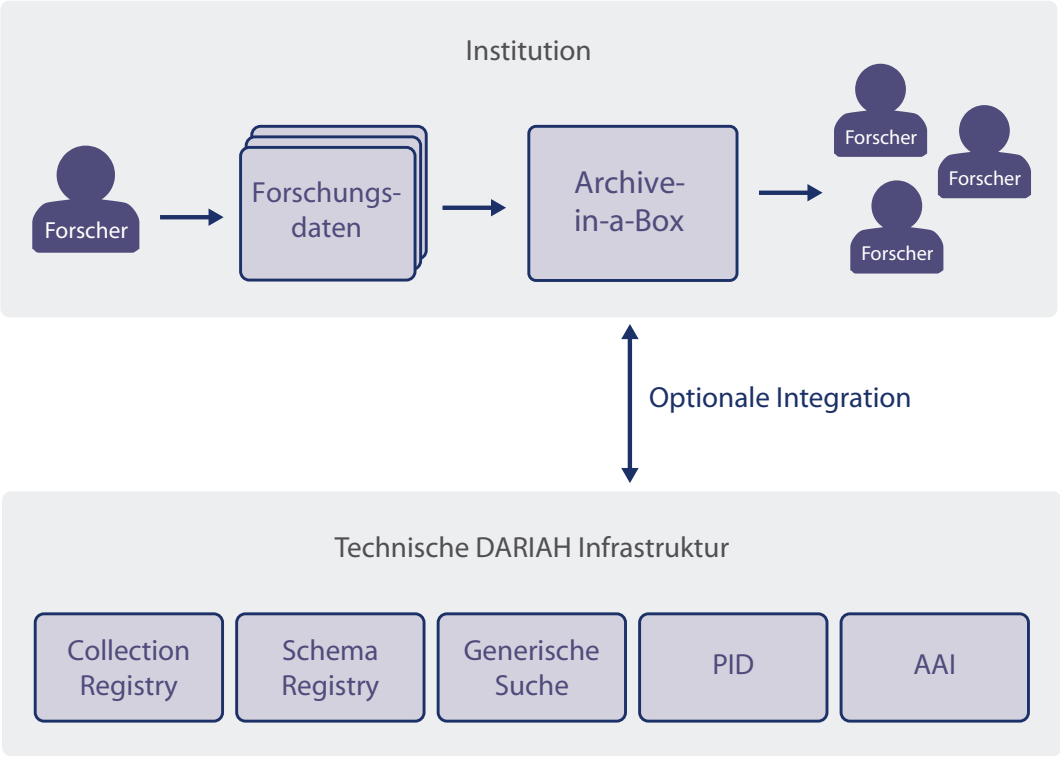


Abb. 6: Archive-in-a-box (Empfehlung).

Das Anlegen und Archivieren einer eigenen Sammlung oder der Transfer bestehender Daten in eine Sammlung wird auf diese Weise ohne besondere Vorkenntnisse möglich.

Bei Fragen zu DARIAH-DE-Empfehlungen wenden Sie sich bitte an
empfehlungen@dariah.eu

DARIAH-DE-Forschungs-infrastruktur

Um Werkzeuge für die Forschungsprozesse der Digital Humanities zur Verfügung stellen zu können, bedarf es einer soliden Forschungsinfrastruktur aus Hard- und Software sowie deren Betreuung, Verwaltung und Überwachung. Diese Forschungsinfrastruktur ist außerdem Grundlage für zukünftig zu entwickelnde Werkzeuge und Virtuelle Forschungsumgebungen. DARIAH-DE sieht eine große Verantwortung vor allem in der Unterstützung von kleineren Forschungsprojekten oder einzelnen Wissenschaftlern, die eine solche Forschungsinfrastruktur nicht selbst aufbauen können. DARIAH-DE fungiert ebenso als Ansprechpartner für größere Verbundforschungsvorhaben und Datenzentren.

DARIAH-DE-Portal – de.dariah.eu

Alle in dieser Broschüre vorgestellten Komponenten sowie weitere Informationen werden im DARIAH-DE-Portal zugänglich und interaktiv miteinander verknüpft sein. Neben der für den Benutzer aufbereiteten Browser-Ansicht werden die zugrunde liegenden Daten aus den *Registries* für die automatische Datenverarbeitung im semantischen Web als *Linked Data* vorliegen. Auf diese Weise können alle Beiträge im Datenpool von DARIAH-EU zusammengespielt werden. Das DARIAH-DE-Portal wird zu einer zentralen Anlaufstelle für die Forschungsinfrastrukturbedürfnisse im digitalen Bereich von Geistes- und Kulturwissenschaftlern ausgebaut.

Hinter den Kulissen (Funktionsweise der Werkzeuge)

Die technische DARIAH-DE-Forschungsinfrastruktur setzt sich aus unabhängig voneinander nutzbaren Softwarekomponenten zusammen, die dem Nutzer zur Verfügung stehen. Beispiele hierfür sind die Demonstratoren oder die Collection Registry. Der Nutzer greift auf diese Softwarekomponenten meist über webbasierte Schnittstellen zu. Im Hintergrund arbeiten weitere Softwarekomponenten, die für den Nutzer weitgehend verborgen bleiben, jedoch Kernaufgaben, wie z. B. die Nutzer-Authentifizierung, übernehmen.

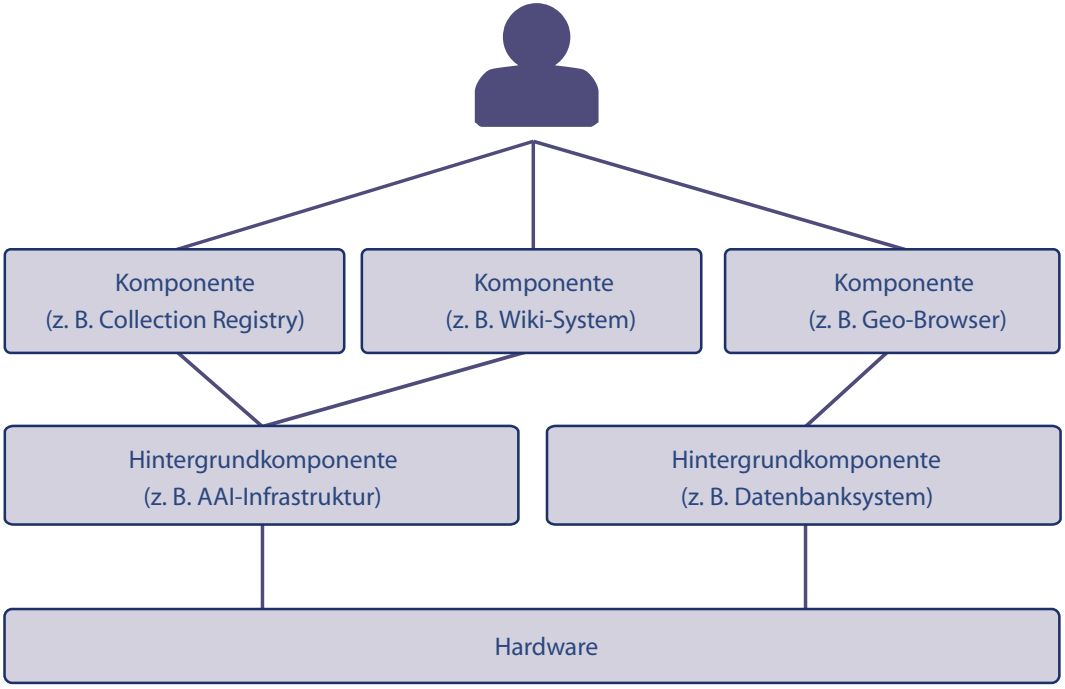


Abb. 7: DARIAH-DE-Forschungsinfrastruktur.

Diese Softwarekomponenten sind auf mehrere Rechenzentren verteilt, wobei die Rechenzentren jeweils unterschiedliche Aufgaben hauptverantwortlich wahrnehmen. Der Vorteil dieser Verteilung liegt in dem optimalen Einsatz von Kernkompetenzen der Partner. Ein einziger zentraler Zugang zu diesen unterschiedlichen Komponenten bringt eine erheblich vereinfachte Nutzung der Forschungsinfrastruktur.

Alle von DARIAH-DE aufgesetzten Forschungsinfrastrukturkomponenten verfügen über eine einheitliche Authentifizierungs- und Autorisierungsinfrastruktur, die es Nutzern erlaubt, mittels eines einzigen Accounts Zugriff auf alle Systeme zu erhalten. Ein umfangreiches Rechte-Management gewährleistet, dass die Nutzer abgesichert in den jeweils eigenen Bereichen tätig sein können. Auf diese Weise werden z. B. die Daten eines Nutzers genauso vor unerlaubtem Zugriff geschützt wie Schemata in der *Schema Registry* oder Quellcode in der Versionsverwaltung.

Langfristige Datenhaltung (Bitstream Preservation, Persistente Identifikatoren, Provenienz)

Während eines Forschungsprojekts fallen Daten unterschiedlicher Art an. Um Forschungsdaten nachweisen sowie einer späteren Nutzung zuführen zu können, müssen sie langfristig aufbewahrt und nachnutzbar gemacht werden. DARIAH-DE entwickelt u. a. eine *Bitstream Preservation*-Komponente, die eine langfristige technische Unversehrtheit der Forschungsdaten garantiert. Außerdem können diese Daten mit Hilfe von persistenten Identifikatoren langfristig und eindeutig identifiziert und damit zitierfähig gemacht werden.

Unter dem Begriff Forschungsdaten versteht DARIAH-DE nicht nur Daten, die während eines Forschungsprozesses entstehen, sondern auch Daten zu den einzelnen Schritten ihrer Erzeugung und Bearbeitung. Dies wird unter dem Begriff Provenienz zusammengefasst. Während der Datenerzeugung werden ausreichend Informationen zu den einzelnen Verarbeitungsschritten aufgezeichnet und mit den Daten zusammen abgelegt, so dass eine ausführliche Dokumentation der Provenienz entsteht. DARIAH-DE erarbeitet Leitfäden und Anwendungsbeispiele, um die Erstellung von Provenienzdaten in Forschungsprojekten nach einheitlichen Standards zu fördern.

Die eigene Datenverarbeitung (Research Software Hosting)

In DARIAH-DE werden unter anderem Lösungen für die Nutzung von Rechenressourcen erprobt. Es wird eine Applikationsumgebung eingerichtet, in der spezialisierte Dienste hochverfügbar angeboten werden können. Nutzer können ihre eigenen Dienste für die geistes- und kulturwissenschaftliche Forschung, sofern diese bestimmten Vorgaben genügen, in die Applikationsumgebung transferieren.

Die Dienste werden, soweit erforderlich, mit den grundlegenden DARIAH-DE-Diensten (wie etwa Authentifizierungs- und Autorisierungsinfrastruktur, Monitoring usw.) verknüpft. Außerdem werden Rechen- und Speicherlösungen der Cloud-Technologie geprüft, um dem Nutzer die vorhandenen Ressourcen kurzfristig, z. B. für eine wissenschaftliche Berechnung, bereitzustellen. Als Fernziel dieser Aktivitäten sind Workspaces für rechen- und speicherintensive Anforderungen vorgesehen, die die Nutzung typischer DARIAH-DE-Dienste und -Anwendungen vereinfachen werden.

Werkzeuge für Programmierer (Developer Portal)

DARIAH-DE bietet eine Umgebung für Softwareentwicklung an. Dieses Entwickler-Portal beinhaltet eine Suite von Standard-Entwickler-Werkzeugen. Jedes Projekt kann on demand seine eigene Entwicklungsumgebung gestalten. Bei DARIAH-DE registrierte Nutzer können an verschiedenen Projekten teilnehmen oder eigene Entwickler-Projekte starten. Für jedes Projekt wird definiert, welche Rollen die verschiedenen Nutzer einnehmen. Die Autorisierung innerhalb der Werkzeuge wird durch die Projektrollen gesteuert. Das öffentliche DARIAH-DE-Entwickler-Portal enthält die folgenden Komponenten:

- Versionsverwaltung mit Subversion
- Ticketing mit Jira
- Wiki mit Confluence
- Automatisierte Integrationstests mit Jenkins
- Quellcodemanagement mit Sonar
- Software Artefakt Management mit Archiva
- Kollaborative Textverarbeitung mit Etherpad

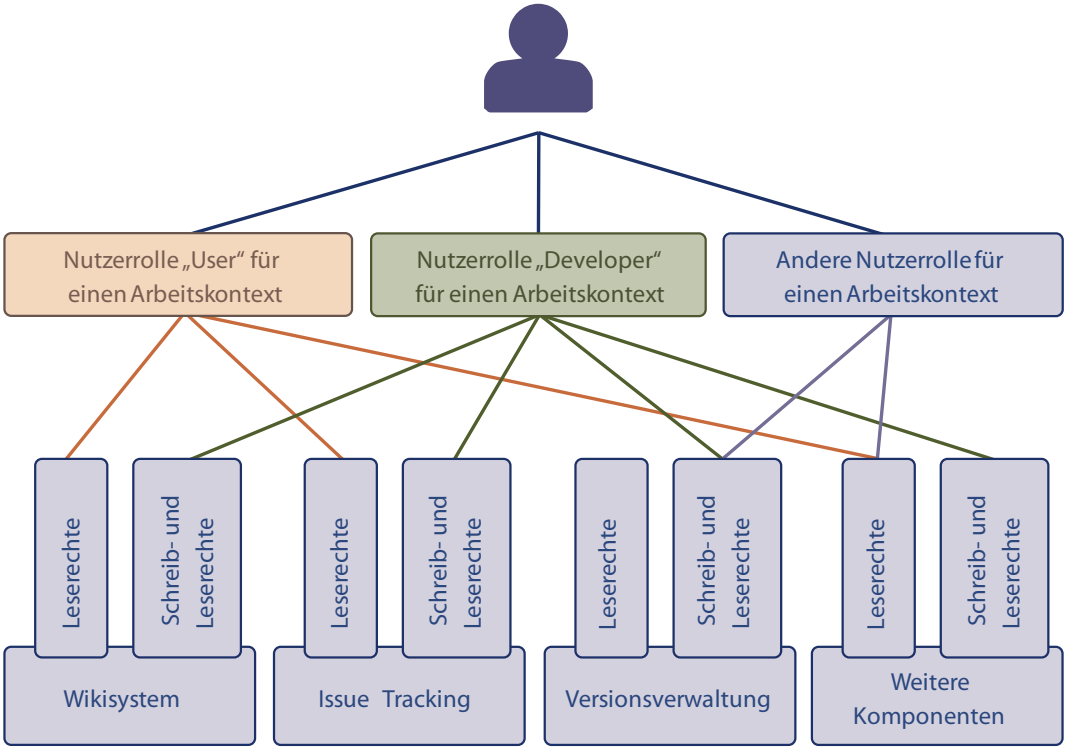


Abb. 8: DARIAH-DE Entwickler-Portal.

Interoperabilität

Die Interoperabilität von Daten- und Metadatenformaten ist sowohl in geistes- und kulturwissenschaftlicher als auch informationstechnologischer Hinsicht ein komplexes Feld und erfordert deshalb die Verwendung von standardisierten Protokollen und Schnittstellen für den maschinellen Datenzugriff. Diese werden in DARIAH-DE angewandt und weiterentwickelt.

Die in DARIAH-DE definierten Empfehlungen und Best Practices als konkrete Handlungsanweisungen für geistes- und kulturwissenschaft-

liche Nutzer gewährleisten die Interoperabilität von Forschungsdaten über Institutions- und Disziplingrenzen hinweg. Hier sind u. a. folgende Fragen von Interesse: Wie können z. B. Historiker auf Daten der Archäologie oder der Judaistik zugreifen? Welche Anforderungen entstehen, wenn Forscher versuchen, auf Daten anderer Institutionen und Disziplinen zuzugreifen, wenn sie sie nachnutzen und in die eigene Forschung einbeziehen wollen?

Die Infrastruktur mitgestalten

Die Schnittstellen aller in DARIAH-DE entwickelten Komponenten der Forschungsinfrastruktur wenden allgemein anerkannte Standards an, so dass die Interoperabilität der Komponenten untereinander sowie mit zukünftig zu entwickelnden oder DARIAH-DE-externen Komponenten sichergestellt ist. Außerdem konzipiert DARIAH-DE Empfehlungen auf technischer Ebene, anhand derer Entwickler eine Orientierung zur Integration neuer Komponenten mit existierenden DARIAH-DE-Komponenten erhalten. Damit ist garantiert, dass DARIAH-DE auf der Basis künftiger spezifischer Anforderungen einzelner Projekte ergänzt und erweitert wird.

Bei Fragen zur Forschungsinfrastruktur von DARIAH-DE wenden Sie sich bitte an

infrastruktur@dariah.eu

Als DARIAH-DE-Nutzer registrieren Sie sich per E-Mail an

register@dariah.eu

DARIAH-DE: Lehre in den Digital Humanities

Trainings und Workshops

DARIAH-DE konzipiert und organisiert Weiterbildungsveranstaltungen, die sich an Experten, Dozenten, Doktoranden sowie an Studierende in der Abschlussphase richten. Dazu gehören Workshops mit methodischen und inhaltlich-thematischen Schwerpunkten.

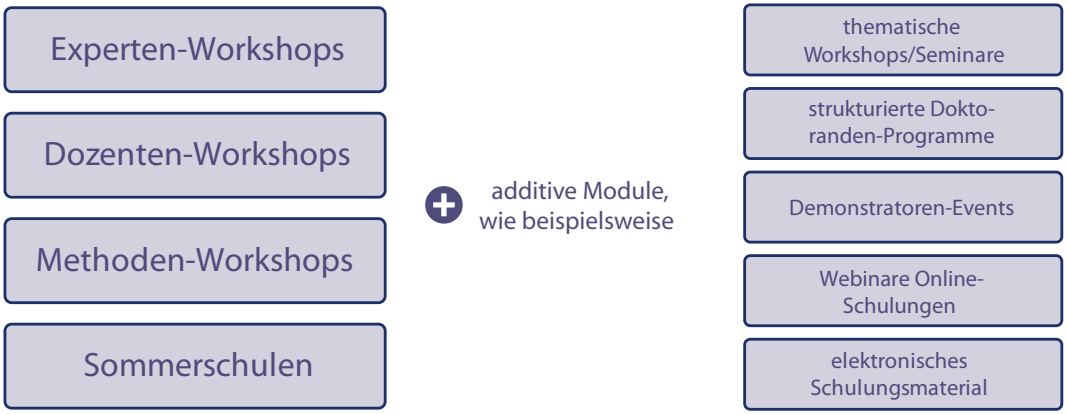


Abb. 9: Veranstaltungsformate und Module der DARIAH-DE-Weiterbildungen.

Vier verschiedene Formate der DARIAH-DE-Weiterbildungsveranstaltungen können durch verschiedene Module ergänzt werden. Diese additiven Module (siehe Abb. 9) können beispielsweise in strukturierte Doktoranden-Programme eingebracht werden, in denen Nachwuchswissenschaftler im Rahmen eines Methoden-Workshops neue Forschungsansätze zur Lösung fachspezifischer Fragestellungen entwickeln. Im Modul der sogenannten „Demonstratoren-Events“ werden anhand von Anwendungsbeispielen u. a. neue Werkzeuge, Dienste oder Virtuelle Forschungsumgebungen vorgestellt und erprobt. Ein Dozenten-Workshop kann beispielsweise mit dem Aufbaumodul für thematische Seminare verknüpft werden, um lösungsorientierte Fragestellungen in der Lehre zu erörtern. Die Workshops können zusätzlich durch elektronisch bereitgestelltes Schulungsmaterial ergänzt bzw. in Form von Webinaren unterstützt werden, die die DARIAH-DE-Angebotspalette abrunden.

Die DARIAH-DE-Trainings und -Workshops werden in Kooperation mit Partnerhochschulen oder -institutionen aus dem Inland und europäischen Ausland ausgerichtet. Die Durchführung beschränkt sich nicht nur auf die DARIAH-DE-Konsortialpartner – vielmehr ist der DARIAH-DE-Projektverbund offen für neue Partnerschaften.

Curricula und DH-Studiengänge

Im Rahmen von DARIAH-DE wurden Informationen zu bestehenden Studienangeboten in den Digital Humanities gesammelt und vermittelt. Eine Broschüre gibt einen Überblick über alle existierenden Programme. Druckexemplare können bei info-cceh@uni-koeln.de angefordert werden.

Eine digitale Fassung ist im Internet abrufbar unter www.cceh.uni-koeln.de/dh-degrees-2011

Eine aktuelle Liste der Studiengänge ist unter www.dig-hum.de/studienstandorte abrufbar.

Abb. 10: Broschüre zu Digital-Humanities-Studiengängen.

DARIAH-DE stimmt sowohl auf nationaler als auch internationaler Ebene die Entwicklung der Curricula in diesen Studiengängen ab.

Darüber hinaus wird eine Digital Humanities-Studie erarbeitet, in der untersucht wird, inwieweit digitale Verfahren bereits Eingang in Lehrveranstaltungen der etablierten geistes- und kulturwissenschaftlichen Disziplinen gefunden haben.

Bei Fragen zu Workshops und Curricula in DARIAH-DE wenden Sie sich bitte an

lehre@dariah.eu

Abkürzungen und URLs

Abkürzungen

ADeX	Archaeological Data eXchange
BMBF	Bundesministerium für Bildung und Forschung
CEI	Charters Encoding Initiative
CIDOC-CRM	CIDOC Conceptual Reference Model
DARIAH-DE	Aufbau von Forschungsinfrastrukturen für die eHumanities
DARIAH-EU	Digital Research Infrastructure for the Arts and Humanities
DH	Digital Humanities
DHd-Blog	Digital Humanities im deutschsprachigen Raum
EpiDoc	Epigraphic Documents
ESFRI	European Strategy Forum on Research Infrastructures Europäisches Strategieforum für Forschungsinfrastrukturen
ERIC	European Research Infrastructure Consortium
GND	Gemeinsame Normdatei
MEI	Music Encoding Initiative
MEISE	Music Encoding Initiative Score Editor
OAI-PMH	Open Archives Initiative – Protocol for Metadata Harvesting
TEI	Text Encoding Initiative
VRE	Virtual Research Environment Virtuelle Forschungsumgebung

URLs

BMBF
www.bmbf.de

Broschüre zu DH-Studiengängen
www.cceh.uni-koeln.de/dh-degrees-2011

DARIAH-Bibliographie
www.zotero.org/groups/dariah_bibliography

DARIAH-DE
www.de.dariah.eu

DARIAH-EU
www.dariah.eu

DHd-Blog
www.dhd-blog.org

ESFRI
ec.europa.eu/research/esfri/

Liste der DH-Studiengänge
www.dig-hum.de/studienstandorte

TextGrid
www.textgrid.de

VRE-Leitfaden der Allianz-Initiative
www.allianzinitiative.de/fileadmin/user_upload/Leitfaden_VRE_de.pdf

Zotero Support
www.zotero.org/support

E-Mailadressen

dhdblog@mpiwg-berlin.de

dariah-sub@sub.uni-goettingen.de

forschung@dariah.eu

forschungsdaten@dariah.eu

empfehlungen@dariah.eu

register@dariah.eu

infrastruktur@dariah.eu

Info-cceh@uni-koeln.de

lehre@dariah.eu

Impressum

Herausgeber

DARIAH-DE
c/o Niedersächsische Staats- und Universitätsbibliothek Göttingen
Abteilung Forschung und Entwicklung
Papendiek 14
37073 Göttingen

Satz/Layout

Heck – Visuelle Kommunikation
Patrick Heck
Thebäerstraße 55
54292 Trier

Druck

Druckerei Schaubs GbR
Gottbillstraße 33 b
D - 54294 Trier

Piktogramme

Font Awesome
fortawesome.github.com/Font-Awesome

Auflage: 5.000

Stand: Juni 2012

Dieses Werk bzw. der Inhalt steht unter einer Creative Commons Namensnennung – Weitergabe unter gleichen Bedingungen 3.0 Deutschland Lizenz.
creativecommons.org/licenses/by-sa/3.0/

DOI: 10.3249/webdoc-3589
resolver.sub.uni-goettingen.de/purl/?webdoc-3589

INSTITUT FÜR EUROPÄISCHE GESCHICHTE

DAASI International

JÜLICH FORSCHUNGSZENTRUM

UNIVERSITÄT PADERBORN
Die Universität der Informationsgesellschaft

KIT Karlsruhe Institute of Technology

SUB NIEDERSÄCHSISCHE STAATS- UND UNIVERSITÄTSBIBLIOTHEK GÖTTINGEN

TECHNISCHE UNIVERSITÄT DARMSTADT

Julius-Maximilians-UNIVERSITÄT WÜRZBURG

GCDH GÖTTINGEN CENTRE FOR DIGITAL HUMANITIES

MAX PLANCK digital library

INSTITUTUM ARCHAEOLOGICUM GERMANICUM

Historisch-Kulturwissenschaftliche Informationsverarbeitung

UNIVERSITY OF BAMBERG
OTTO-FRIEDRICH-UNIVERSITÄT BAMBERG

berlin-brandenburgische AKADEMIE DER WISSENSCHAFTEN

GWDG

rzg
IPP RECHEN-ZENTRUM GARCHING

SALOMON LUDWIG STEINHEIM INSTITUT FÜR DEUTSCH-JÜDISCHE GESCHICHTE

DARIAH-DE erforscht und entwickelt eine digitale Forschungsinfrastruktur für die Geistes- und Kulturwissenschaften. Dazu werden Curricula neu entworfen, elektronisch unterstützte Forschungsmethoden und Verfahren evaluiert und entwickelt, Konzepte für den Umgang mit Forschungsdaten erarbeitet sowie die technische Basisinfrastruktur bereitgestellt.

www.de.dariah.eu